



Geometry-Aware Neighborhood Search for Learning Local Models for Image Superresolution

Julio Cesar Ferreira, Elif Vural, Christine Guillemot

► To cite this version:

Julio Cesar Ferreira, Elif Vural, Christine Guillemot. Geometry-Aware Neighborhood Search for Learning Local Models for Image Superresolution. IEEE Transactions on Image Processing, 2016, 25 (3), pp.14. 10.1109/tip.2016.2522303 . hal-01388955

HAL Id: hal-01388955

<https://inria.hal.science/hal-01388955>

Submitted on 15 Nov 2016

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Geometry-Aware Neighborhood Search for Learning Local Models for Image Super-Resolution

Julio Cesar Ferreira, Elif Vural, and Christine Guillemot

Abstract—Local learning of sparse image models has proven to be very effective to solve inverse problems in many computer vision applications. To learn such models, the data samples are often clustered using the K-means algorithm with the Euclidean distance as a dissimilarity metric. However, the Euclidean distance may not always be a good dissimilarity measure for comparing data samples lying on a manifold. In this paper, we propose two algorithms for determining a local subset of training samples from which a good local model can be computed for reconstructing a given input test sample, where we take into account the underlying geometry of the data. The first algorithm, called Adaptive Geometry-driven Nearest Neighbor search (AGNN), is an adaptive scheme which can be seen as an out-of-sample extension of the replicator graph clustering method for local model learning. The second method, called Geometry-driven Overlapping Clusters (GOC), is a less complex nonadaptive alternative for training subset selection. The proposed AGNN and GOC methods are evaluated in image super-resolution and shown to outperform spectral clustering, soft clustering, and geodesic distance based subset selection in most settings.

Index Terms—Clustering, patch manifolds, nearest neighbor search, image super-resolution, image restoration.

I. INTRODUCTION

THE super-resolution problem can be formulated as a linear inverse problem by modeling the image deformation via a linear system. This problem is generally ill-posed and the solutions often rely on some a priori information about the image to be reconstructed. Research in the recent years has proven that adopting an appropriate sparse image model can yield quite satisfactory reconstruction qualities. Sparse representations are now used to solve inverse problems in many computer vision applications, such as super-resolution [1], [2], [3], [4]; denoising [1], [5], [6]; compressive sensing [7], [8], [9]; and deblurring [1], [2]. While several works assume that the image to be reconstructed has a sparse representation in a large overcomplete dictionary [4], [5], it has also been observed that representing the data with small, local models (such as subspaces) might have benefits over a single and global model since local models may be more adaptive and capture better the local variations in data characteristics [1], [2], [10]. The image restoration methods in [1] and [2] propose a patch-based processing of images, where the training patches are first clustered and then a principal component

analysis (PCA) basis is learned in each cluster. The idea of learning adaptive models from groups of similar patches for image restoration has been exploited in several recent works [11], [12], [13].

When learning local models, the assessment of the similarity between image patches is of essential importance. Different similarity measures lead to different partitionings of data, which may eventually change the learned models significantly. Many algorithms constructing local models assess similarity based on the Euclidean distance between samples. For example in [1] and [2] image patches are clustered using the K-means algorithm, where patches having a small Euclidean distance are grouped together to learn a PCA basis. Test patches are then reconstructed under the assumption that they are sparsely representable in this basis.

However, patches sampled from natural images are highly structured and constitute a low-dimensional subset of the high-dimensional ambient space. In fact, natural image patches are commonly assumed to lie close to a low-dimensional manifold [14], [15]. Similarly, in the deconvolution method proposed in [10], image patches are assumed to lie on a large patch manifold, which is decomposed into a collection of locally linear models learned by clustering and computing local PCA bases. The geometric structure of a patch manifold depends very much on the characteristics of the patches constituting it; the manifold is quite nonlinear especially in regions where patches have a rich texture. When evaluating the similarity between patches on a patch manifold, care should be taken especially in high-curvature regions, where Euclidean distance loses its reliability as a dissimilarity measure. In other words, in the K-means based setting of [1] and [2], one may obtain a good performance only if the local PCA basis agrees with the local geometry of the patch manifold, i.e., the most significant principal directions should correspond to the tangent directions on the patch manifold so that data can be well approximated with a sparse linear combination of only a few basis vectors. While this easily holds in low-curvature regions of the manifold where the manifold is flat, in high-curvature regions, the subspace spanned by the most significant principal vectors computed from the nearest Euclidean-distance neighbors of a reference point may diverge significantly from the tangent space of the manifold if the neighborhood size is not selected properly [16], [17]. This is illustrated in Figure 1, where the first few significant principal directions fail to approximate the tangent space because the manifold bends over itself as in Figure 1(b), or because the curvature principal components dominate the tangential principal components as in Figure 1(c).

In this work, we focus on the image super-resolution

J. C. Ferreira is with the Goiano Federal Institute of Education, Science and Technology, Urutai, 35790-000 Brazil. e-mail: julio.ferreira@ifgoiano.edu.br

E. Vural is with Middle East Technical University, Ankara, 06800 Turkey. e-mail: velif@metu.edu.tr

C. Guillemot is with INRIA, Rennes, 35000 France. e-mail: christine.guillemot@inria.fr

Most part of the work was performed while the first two authors were in INRIA.

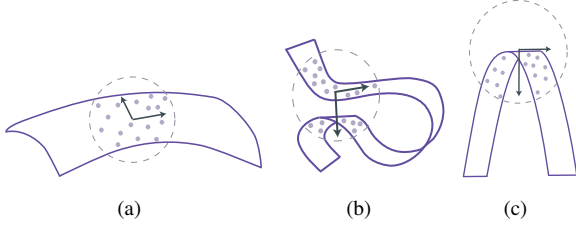


Fig. 1: PCA basis vectors computed with data sampled from a neighborhood on a manifold. In (a), the two most significant principal directions correspond to tangent directions and PCA computes a local model coherent with the manifold geometry. In (b), PCA fails to recover the tangent space as the manifold bends over itself and the neighborhood size is not selected properly. In (c), as the curvature component is stronger than the tangential components, the subspace spanned by the two most significant PCA basis vectors again fails to approximate the tangent space.

inverse problem and consider approaches based on sparse representations of images in locally learned subspaces. We present geometry-driven strategies to select subsets of data samples for learning local models. Given a test sample, we address the problem of determining a local subset of the training samples, i.e., a neighborhood of the test sample, from which a good local model can be computed for reconstructing the test sample, where we take into account the underlying geometry of the data. Hence, the idea underlying this work is to compute local models that agree with the low-dimensional intrinsic geometry of data. Low dimensionality allows sparse representations of data, and the knowledge of sparsity can be efficiently used for solving inverse problems in image restoration.

Training subsets for learning local models can be determined in two ways; adaptively or nonadaptively. In adaptive neighborhood selection, a new subset is formed on the fly for each test sample, whereas in nonadaptive neighborhood selection one subset is chosen for each test sample among a collection of training subsets determined beforehand in a learning phase. Adaptive selection has the advantage of flexibility, as the subset formed for a particular test sample fits its characteristics better than a predetermined subset, but the drawback is the higher complexity. In this work, we study both the adaptive and the nonadaptive settings and propose two different algorithms for geometry-aware local neighborhood selection.

We first present an adaptive scheme, called Adaptive Geometry-driven Nearest Neighbor search (AGNN). Our method is inspired by the Replicator Graph Clustering (RGC) [18] algorithm and can be regarded as an out-of-sample extension of RGC for local model learning. Given a test sample, the AGNN method computes a diffused affinity measure between each test sample and the training samples in a manner that is coherent with the overall topology of the data graph. The nearest neighbor set is then formed by selecting the training samples that have the highest diffused affinities with the test sample.

The evaluation of the adaptive AGNN method in super-resolution experiments shows a quite satisfactory image reconstruction quality. We then propose a nonadaptive scheme called Geometry-driven Overlapping Clusters (GOC), which seeks a less complex alternative for training subset selection.

The method computes a collection of training subsets in a prior learning phase in the form of overlapping clusters. The overlapping clusters are formed by first initializing the cluster centers and then expanding each cluster around its central sample by following the K -nearest neighborhood connections on the data graph. What really determines the performance of the GOC method is the structure of the clusters, driven by the number of neighbors K and the amount of expansion. We propose a geometry-based strategy to set these parameters, by studying the rate of decay of PCA coefficients of data samples in the cluster, thereby characterizing how close the cluster lies to a low-dimensional subspace.

Note that, while the proposed AGNN and GOC algorithms employ similar ideas to those in manifold clustering methods, our study differs from manifold clustering as we do not aim to obtain a partitioning of data. Instead, given a test sample to be reconstructed, we focus on the selection of a local subset of training data to learn a good local model. We evaluate the performance of our methods in image super-resolution application. The results show that the proposed similarity assessment strategies can provide performance gains compared to the Euclidean distance, especially for superresolving images with rich texture where patch manifolds are highly nonlinear. When applying the proposed method in the super-resolution problem, we select the NCSR algorithm [1] as a reference method, which currently leads the state of the art in super-resolution. We first show that the proposed AGNN and GOC methods outperform reference subset selection strategies such as spectral clustering, soft clustering, and geodesic distance based neighborhood selection. Finally, we perform comparative experiments with the NCSR [1], ASDS [2], and SPSR [19] super-resolution algorithms, which suggest that the proposed methods can be successfully applied in super-resolution for taking the state of the art one step further. We also discuss the applicability of the presented methods to other image restoration problems such as deblurring and denoising.

The rest of the paper is organized as follows. In Section II we give an overview of manifold-based clustering methods. In Section III we formulate the neighborhood selection problem studied in this paper. In Section IV we discuss the proposed AGNN method. Then in Section V we describe the GOC algorithm. In Section VI we present experimental results, and in Section VII we conclude.

II. CLUSTERING ON MANIFOLDS: RELATED WORK

As our study has close links with the clustering of low-dimensional data, we now give a brief overview of some clustering methods for data on manifolds. The RGC method [18], from which the proposed AGNN method has been inspired, first constructs a data graph. An initial affinity matrix is then computed based on the pairwise similarities between data samples. The affinity matrix is iteratively updated such that the affinities between all sample pairs converge to the collective affinities that consider all paths on the data graph. Spectral clustering is another well-known algorithm for graph-based clustering [20], [21]. Samples are clustered with respect to a low-dimensional embedding given by the functions of

slowest variation on the data graph, which encourages assigning neighboring samples with strong edge weights to the same cluster. The Laplacian eigenmaps method [22] builds on the same principle; however, it targets dimensionality reduction.

Geodesic clustering provides an extension of the K -means algorithm to cluster data lying on a manifold, where the Euclidean distance is replaced with the geodesic distance [23], [24]. In [25], a method is proposed for clustering data lying on a manifold, which extends the graph-based semi-supervised learning algorithm in [26] to a setting with unlabeled data. The diffusion matrix that diffuses known class labels to unlabeled data in [26] is interpreted as a diffusion kernel in [25], which is then used for determining the similarity between data samples to obtain clusters. The works in [27], [28] also use the geodesic distance as a dissimilarity measure. They propose methods for embedding the manifold into the tangent spaces of some selected reference points and perform a fast approximate nearest neighbor search on the space of embedding.

While the above algorithms consider all data samples to lie on a single manifold, several other methods model low-dimensional data as samples from multiple manifolds and study the determination of these manifolds. An expectation maximization approach is employed in [29] to partition the data into manifolds. The points on each manifold are then embedded into a lower-dimensional domain. The method in [30] computes a sparse representation of each data sample in terms of other samples, where high coefficients are encouraged for nearby samples. Once the sparse coefficients are computed, data is grouped into manifolds simply with spectral clustering. The method in [31] extends several popular nonlinear dimensionality reduction algorithms to the Riemannian setting by replacing the Euclidean distance with the Riemannian distance. It is then shown that, if most data connections lie within the manifolds rather than between them, the proposed Riemannian extensions yield clusters corresponding to different manifolds.

Finally, the generation of overlapping clusters in GOC is also linked to soft clustering [32]. Rather than strictly partitioning the data into a set of disjoint groups, a membership score is computed between each data sample and each cluster center in soft clustering. The cluster centers are then updated by weighing the samples according to the membership scores. In [33], a manifold extension of soft clustering is proposed, where the membership scores are computed with a geodesic kernel instead of the Euclidean distance.

III. RATIONALE AND PROBLEM FORMULATION

In patch-based image processing, one often would like to develop tools that can capture the common structures inherently present in patches and use this information for the efficient treatment of images. One important example is the invariance to geometric transformations. In practical image formation scenarios, different regions of the image are likely to observe the same structure, exposed, however, to different geometric transformations in different parts of the image plane. While most patch-based methods inherently achieve invariance to translations as they extract patches from the image over sliding windows, more complex transformations

such as rotations and scale changes are more difficult to handle in evaluating the structural similarities between patches. In addition to geometric transformation models, structural similarities between image patches may be stemming from many other low-dimensional, possibly parametrizable patch models as well. In [15], several parametrizable patch manifold models are explored such as oscillating textures and cartoon images. In the treatment or reconstruction of image patches, local models computed from patches sharing the same structure reflect the local geometry of the patch manifold, while the comparison of patch similarities based on Euclidean distance does not necessarily achieve this. In this paper, we propose similarity assessment strategies that better take structural similarities into account than the simple Euclidean distance in image restoration.

Given observed measurements \mathbf{y} , the ill-posed inverse problem can be generally formulated in a Banach space as

$$\mathbf{y} = \Theta \mathbf{x} + \nu \quad (1)$$

where Θ is a bounded operator, \mathbf{x} is an unknown data point and ν is an error term. In image restoration \mathbf{y} is the vectorized form of an observed image, Θ is a degradation matrix, \mathbf{x} is the vectorized form of the original image, and ν is an additive noise vector. There are infinitely many possible data points \mathbf{x} that explain \mathbf{y} ; however, image restoration algorithms aim to reconstruct the original image \mathbf{x} from the given measurements \mathbf{y} , often by using some additional assumptions on \mathbf{x} . When Θ is composed of a down-sampling operator and a blurring operator, image restoration becomes single image super-resolution.

In image restoration with sparse representations, \mathbf{x} can be estimated by minimizing the cost function

$$\hat{\alpha} = \arg \min_{\alpha} \left\{ \|\mathbf{y} - \Theta \Phi \alpha\|_2^2 + \lambda \|\alpha\|_1 \right\} \quad (2)$$

where Φ is a dictionary, α is the sparse representation of \mathbf{x} in Φ , and $\lambda > 0$ is a regularization parameter. It is common to reconstruct images patch by patch and model the patches of \mathbf{x} as sparsely representable in Φ . Representing the extraction of the j -th patch x_j of \mathbf{x} with a matrix multiplication as $x_j = R_j \mathbf{x}$, the reconstruction of the overall image \mathbf{x} can be represented via the operator \circ as shown in [1], [2]. If the dictionary Φ is well-chosen, one can efficiently model the data points \mathbf{x} using their sparse representations in Φ . Once the sparse coefficient vector α is estimated, one can reconstruct the image \mathbf{x} as

$$\hat{\mathbf{x}} = \Phi \circ \hat{\alpha}. \quad (3)$$

While a global model is considered in the above problem, several works such as [1], [2], [34] propose to reconstruct image patches based on sparse representations in local models. In this case, one aims to reconstruct the j -th patch x_j of the unknown image \mathbf{x} from its degraded observation y_j by selecting a local model that is suitable for y_j . The problem in (2) is then reformulated as

$$\hat{\alpha}_j = \arg \min_{\alpha_j} \left\{ \|y_j - \Theta \Phi_j \alpha_j\|_2^2 + \lambda \|\alpha_j\|_1 \right\} \quad (4)$$

where y_j is the j -th patch from the observed image \mathbf{y} , Φ_j is a local (PCA) basis chosen for the reconstruction of y_j , and

$\hat{\alpha}_j$ is the coefficient vector. The unknown patch x_j is then reconstructed as $\hat{x}_j = \Phi_j \hat{\alpha}_j$. The optimization problem in (4) forces the coefficient vector $\hat{\alpha}_j$ to be sparse. Therefore, the accuracy of the reconstructed patch \hat{x}_j in approximating the unknown patch x_j depends on the reliability of the local basis Φ_j , i.e., whether signals are indeed sparsely representable in Φ_j .

The main idea proposed in this paper is to take into account the manifold structure underlying the data when choosing a neighborhood of training data points to learn a local basis. Our purpose is to develop a dissimilarity measure that is better suited to the local geometry of the data than the Euclidean distance and also to make the neighborhood selection procedure as adaptive as possible to the test samples to be reconstructed.

Let $\mathcal{D} = \{d_i\}_{i=1}^m$ be a set of m training data points $d_i \in \mathbb{R}^n$ lying on a manifold \mathcal{M} and let $\mathcal{Y} = \{y_j\}_{j=1}^M$ be a set of M test data points $y_j \in \mathbb{R}^n$. As for the image restoration problem in (4), each test data point y_j corresponds to a degraded image patch, and the training data points in \mathcal{D} are used to learn the local bases Φ_j . The test samples y_j are not expected to lie on the patch manifold \mathcal{M} formed by the training samples; however, one can assume y_j to be close to \mathcal{M} unless the image degradation is very severe.

We then study the following problem. Given an observation $y_j \in \mathcal{Y}$ of an unknown image patch x_j , we would like to select a subset $S \subset \mathcal{D}$ of training samples such that the PCA basis Φ_j computed from S minimizes the reconstruction error $\|x_j - \hat{x}_j\|$, where the unknown patch x_j is reconstructed as $\hat{x}_j = \Phi_j \hat{\alpha}_j$, and the sparse coefficient vector is given by

$$\hat{\alpha}_j = \arg \min_{\alpha_j} \left\{ \|y_j - \Theta \Phi_j \alpha_j\|_2^2 + \lambda \|\alpha_j\|_1 \right\}. \quad (5)$$

Since the nondeformed sample x_j is not known, it is clearly not possible to solve this problem directly. In this work, we propose some constructive solutions to guide the selection of S by assuming that y_j lies close to \mathcal{M} . As the manifold \mathcal{M} is not known analytically, we capture the manifold structure of training data \mathcal{D} by building a similarity graph whose nodes and edges represent the data points and the affinities between them. In Sections IV and V we describe the AGNN and the GOC methods, which respectively propose an adaptive and a nonadaptive solution for training subset selection for local basis learning from the similarity graph.

IV. ADAPTIVE GEOMETRY-DRIVEN NEAREST NEIGHBOR SEARCH

In this section, we present the Adaptive Geometry-driven Nearest Neighbor search (AGNN) strategy for selecting the nearest neighbors of each test data point within the training data points with respect to an intrinsic manifold structure. Our subset selection method builds on the RGC algorithm [18], which targets the clustering of data with respect to the underlying manifold. The RGC method seeks a globally consistent affinity matrix that is the same as its diffused version with respect to the underlying graph topology. However, the RGC method focuses only on the initially available training samples and does not provide a means of handling initially unavailable test samples. We thus present an out-of-sample

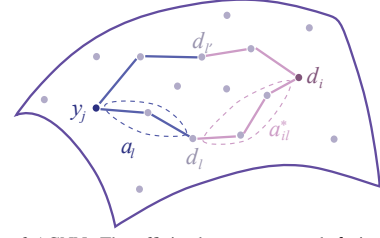


Fig. 2: Illustration of AGNN. The affinity between y_j and d_l is a_l , and the affinity between d_l and d_i is a_{il} . The intermediate node d_l contributes by the product $a_l a_{il}$ to the overall affinity between y_j and d_i . The sample $d_{l'}$ is just another intermediate node like d_l . Summing the affinities via all possible intermediate nodes (i.e., all training samples), the overall affinity is obtained as in (9).

generalization of RGC and propose a strategy to compute and diffuse the affinities between the test sample and all training samples in a way that is consistent with the data manifold.

In the RGC algorithm, given a set of data points \mathcal{D} , an affinity matrix $A = (a_{il})$ is first computed. The elements a_{il} of A measure the similarity between the data points d_i and d_l . A common similarity measure is the Gaussian kernel

$$a_{il} = \exp \left(-\frac{\|d_i - d_l\|^2}{nc_1^2} \right) \quad (6)$$

where $\|\cdot\|$ denotes the ℓ_2 -norm on \mathbb{R}^n and c_1 is a constant. Then, the initial affinities are updated with respect to the underlying manifold as follows. The affinities are diffused by looking for an A matrix such that each row A_i of A maximizes

$$A_i^T = \arg \max_v (v^T A v). \quad (7)$$

Since the maximization problem on the right hand side of (7) is solved by an eigenvector of A , the method seeks an affinity matrix such that the similarities between the data sample d_i and all the other samples in \mathcal{D} (given by the row A_i) are proportional to the diffused version of the similarities in A_i over the whole manifold via the product AA_i^T ; i.e., an affinity matrix is searched such that $A_i^T \propto AA_i^T$. The optimization problem in (7) is solved with an iterative procedure based on a game theoretical approach to obtain a diffused affinity matrix A^* . The diffusion of the affinities are constrained to the s nearest neighbors of each point d_i .

In our AGNN method, we first compute and diffuse the affinities of training samples in \mathcal{D} as proposed in [18]. This gives us a similarity measure coherent with the global geometry of the manifold. Meanwhile, unlike in RGC, our main purpose is to select a subset $S \subset \mathcal{D}$ of training samples for a given test sample $y_j \in \mathcal{Y}$. We thus need a tool for generalizing the above approach for test samples.

We propose to compute the affinities between y_j and \mathcal{D} by employing A^* as follows. Given a test data point $y_j \in \mathcal{Y}$, we first compute an initial affinity vector a whose i -th entry

$$a_i = \exp \left(-\frac{\|y_j - d_i\|^2}{nc_1^2} \right) \quad (8)$$

measures the similarity between y_j and the training sample d_i . We then update the affinity vector as follows. Denoting the entries of the diffused affinity matrix A^* by a_{il}^* , first the product $a_{il}^* a_l$ should give the component of the overall affinity

Algorithm 1 Adaptive Geometry-driven Nearest Neighbor search (AGNN)

1: **Input:**
 $\mathcal{D} = \{d_i\}_{i=1}^m$: Set of training samples
 $y_j \in \mathcal{Y}$: Test sample
 c_1, c_2, κ : Algorithm parameters

2: **AGNN Algorithm:**
 3: Form affinity matrix A of training samples with respect to (6).
 4: Diffuse the affinities in A to obtain A^* as proposed in the RGC method [18].
 5: Initialize the affinity vector a between test sample y_j and the training samples as in (8).
 6: Diffuse the affinities in a to obtain a^* with respect to (10).
 7: Determine set S of nearest neighbors of y_j by selecting the training samples with the highest affinities as in (11).
 8: **Output:**
 S : Set of nearest neighbors of y_j in \mathcal{D} .

between y_j and d_i that is obtained through the sample d_l : if there is a sample d_l that has a high affinity with both d_i and y_j , this means that the affinity between d_i and y_j should also be high due to the connection established via the intermediate node d_l (see the illustration in Figure 2). Note that the formulation in (7) also relies on the same idea. We thus update the affinity vector a such that its i -th entry a_i becomes proportional to

$$\sum_{l=1}^m a_{il}^* a_l \quad (9)$$

i.e., the total affinity between samples d_i and y_j obtained through all nodes d_l in the training data graph. This suggests that the initial affinities in the vector a should be updated as A^*a , which corresponds to the diffusion of the affinities on the graph. Repeating this diffusion process κ times, we get the diffused affinities of the test sample as

$$a^* = (A^*)^\kappa a \quad (10)$$

where a_i^* gives the final diffused affinity between y_j and d_i . This generalizes the idea in (7) to initially unavailable data samples; and hence, provides an out-of-sample extension of the diffusion approach in RGC. The parameter κ should be chosen in a way to permit a sufficient diffusion of the affinities. However, it should not be too large in order not to diverge too much from the initial affinities in a . In our experiments we have observed that $\kappa = 2$ gives good results in general.

Once the affinities a^* are computed, the subset S consisting of the nearest neighbors of y_j can be obtained as the samples in \mathcal{D} whose affinities to y_j are higher than a threshold

$$S = \{d_i \in \mathcal{D} : a_i^* \geq c_2 \max_l a_l^*\} \quad (11)$$

where $0 < c_2 < 1$. The samples in S are then used for learning a PCA basis to reconstruct y_j . The threshold c_2 should be chosen sufficiently high to select only the similar patches to the reference patch, however, it should not be selected too high in order to have sufficiently many neighbors necessary for computing a basis. If S contains too few samples, the threshold c_2 can be adapted to increase the number of samples or a sufficient number of points with highest affinities can be directly included in S . The proposed AGNN method for determining training subsets gets around the problem depicted in Figure 1(b), since points lying at different sides of a

manifold twisting onto itself have a small diffused affinity and are not included in the same subset. A summary of the proposed AGNN method is given in Algorithm 1.

V. GEOMETRY-DRIVEN OVERLAPPING CLUSTERS

As we will see in Section VI, the AGNN method presented in Section IV is efficient in terms of image super-resolution performance. However, it may have a high computational complexity and considerable memory requirements in settings with a large training set \mathcal{D} , as the size of the affinity matrix grows quadratically with the number of training samples and the subset selection is adaptive (repeated for each test sample). For this reason, we propose in this section the Geometry-driven Overlapping Clusters (GOC) method, which provides a computationally less complex solution for obtaining the nearest neighbors of test samples.

The GOC algorithm computes a collection $\{S_k\}_{k=1}^C$ of subsets $S_k \subset \mathcal{D}$ of the training data set, which are to be used in local basis computation. Contrary to the AGNN method, the subsets $S_k \subset \mathcal{D}$ are determined only using the training data and are not adapted to the test samples. However, the number C of subsets should then be sufficiently large to have the desired adaptivity for capturing arbitrary local variations. Due to the large number of subsets, S_k are not disjoint in general; hence, can be regarded as overlapping clusters. In the following, we first describe our method for forming the clusters and then propose a strategy to select some parameters that determine the size and the structure of the clusters.

Given the number of clusters C to be formed, we first determine the central data point $\mu_k \in \mathcal{D}$ of each cluster S_k . In our implementation, we achieve this by first clustering \mathcal{D} with the K-means algorithm, and then choosing each μ_k as the point in \mathcal{D} that has the smallest Euclidean distance to the center of the k -th cluster given by K-means.

The training data points μ_k are used as the kickoff for the formation of the clusters S_k . Given the central sample μ_k , the cluster S_k is formed iteratively with the GOC algorithm illustrated in Figure 3 as follows. We first initialize S_k as

$$S_k^0 = \mathcal{N}_K(\mu_k) \quad (12)$$

where $\mathcal{N}_K(\mu_k)$ denotes the set of the K -nearest neighbors of μ_k in \mathcal{D} with respect to the Euclidean distance. Then in each iteration l , we update the cluster S_k^l as

$$S_k^l = S_k^{l-1} \cup \bigcup_{d_i \in S_k^{l-1}} \mathcal{N}_K(d_i) \quad (13)$$

by including all samples in the previous iteration as well as their K -nearest neighbors. Hence, the clusters are gradually expanded by following the nearest neighborhood connections on the data graph. This procedure is repeated for L iterations so that the final set of clusters is given by

$$\{S_k\}_{k=1}^C = \{S_k^L\}_{k=1}^C. \quad (14)$$

The expansion of the clusters is in a similar spirit to the affinity diffusion principle of AGNN; however, is computationally much less complex.

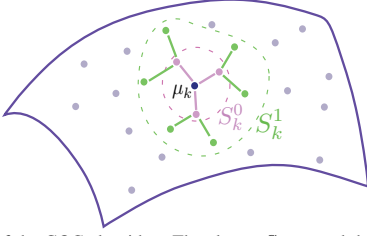


Fig. 3: Illustration of the GOC algorithm. The cluster S_k around the central sample μ_k is formed gradually. S_k is initialized with S_k^0 containing the K nearest neighbors of μ_k ($K = 3$ in the illustration). Then in each iteration l , S_k^l is expanded by adding the nearest neighbors of recently added samples.

In the simple strategy presented in this section, we have two important parameters to set, which essentially influence the performance of learning: the number of iterations L and the number of samples K in each small neighborhood. In the following, we propose an algorithm to adaptively set these parameters based on the local geometry of data. Our method is based on the observation that the samples in each cluster will eventually be used to learn a local subspace that provides an approximation of the local tangent space of the manifold. Therefore, S_k should lie close to a low-dimensional subspace in \mathbb{R}^n , so that nearby test samples can be assumed to have a sparse representation in the basis Φ_k computed from S_k . We characterize the concentration of the samples in S_k around a low-dimensional subspace by the decay of the coefficients of the samples in the local PCA basis.

We omit the cluster index k for a moment to simplify the notation and consider the formation of a certain cluster $S = S_k$. With a slight abuse of notation, let $S^{L,K}$ stand for the cluster S that is computed by the algorithm described above with parameters L and K . Let $\Phi = [\phi_1 \dots \phi_n]$ be the PCA basis computed with the samples in S , where the principal vectors $\phi_1, \dots, \phi_n \in \mathbb{R}^n$ are sorted with respect to the decreasing order of the absolute values of their corresponding eigenvalues. For a training point $d_i \in S$, let $\bar{d}_i = d_i - \eta_S$ denote the shifted version of d_i where $\eta_S = |S|^{-1} \sum_{d_i \in S} d_i$ is the centroid of cluster S . We define

$$I(L, K) = \min \left\{ \iota \mid \sum_{q=1}^L \sum_{d_i \in S^{L,K}} \langle \phi_q, \bar{d}_i \rangle^2 \right. \\ \left. \geq c_3 \sum_{q=1}^n \sum_{d_i \in S^{L,K}} \langle \phi_q, \bar{d}_i \rangle^2 \right\} \quad (15)$$

which gives the smallest number of principal vectors to generate a subspace that captures a given proportion c_3 of the total energy of the samples in S , where $0 < c_3 < 1$. We propose to set the parameters L, K by minimizing the function $I(L, K)$, which gives a measure of the concentration of the energy of S around a low-dimensional subspace. However, in the case that S contains $m \leq n$ samples where n is the dimension of the ambient space, the subspace spanned by the first $m - 1$ principal vectors always captures all of the energy in S ; therefore $I(L, K)$ takes a relatively small value; i.e., $I(L, K) \leq m - 1$. In order not to bias the algorithm towards reducing the size of the clusters as a result of this, a

Algorithm 2 Geometry-driven Overlapping Clusters (GOC)

1: **Input:**
 $\mathcal{D} = \{d_i\}_{i=1}^m$: Set of training samples
 C : Number of clusters
 c_3 : Algorithm parameter

2: **GOC Algorithm:**
3: Determine cluster centers μ_k of all C clusters (possibly with the K-means algorithm).
4: **for** $k = 1, \dots, C$ **do**
5: Fix parameter $L' = L_0$ at an initial value L_0 .
6: **for** $K' = 1, \dots, K_{max}$ **do**
7: Form cluster $S_k = S^{L_0, K'}$ as described in (12)-(14).
8: Evaluate decay rate function $\tilde{I}(L_0, K')$ given in (16).
9: **end for**
10: Set K as the K' value that minimizes $\tilde{I}(L_0, K')$.
11: **for** $L' = 1, \dots, L_{max}$ **do**
12: Form cluster $S_k = S^{L', K}$ as described in (12)-(14).
13: Evaluate decay rate function $\tilde{I}(L', K)$ given by (16).
14: **end for**
15: Set L as the L' value that minimizes $\tilde{I}(L', K)$.
16: Determine cluster S_k as $S^{L, K}$ with the optimized parameters.
17: **end for**
18: **Output:**
 $\{S_k\}_{k=1}^C$: Set of overlapping clusters in \mathcal{D} .

normalization of the function $I(L, K)$ is required. We define

$$\tilde{I}(L, K) = I(L, K) / \min\{|S^{L,K}| - 1, n\} \quad (16)$$

where $|\cdot|$ denotes the cardinality of a set. The denominator $\min\{|S^{L,K}| - 1, n\}$ of the above expression gives the maximum possible value of $I(L, K)$ in cluster $S^{L,K}$. Hence, the normalization of the coefficient decay function by its maximum value prevents the bias towards small clusters.

We can finally formulate the selection of L, K as $(L, K) = \arg \min_{(L', K') \in \Lambda} \tilde{I}(L', K')$ where Λ is a bounded parameter domain. This optimization problem is not easy to solve exactly. One can possibly evaluate the values of $\tilde{I}(L, K)$ on a two-dimensional grid in the parameter domain. However, in order to reduce the computation cost, we approximately minimize the objective by optimizing one of the parameters and fixing the other in each iteration. We first fix the number of iterations L at an initial value and optimize the number of neighbors K . Then, updating and fixing K , we optimize L .

The computation of the parameters L and K with the above procedure determines the clusters as in (14). The samples in each cluster S_k are then used for computing a local basis Φ_k . The proposed GOC method is summarized in Algorithm 2. Since the proposed GOC method determines the clusters not only with respect to the connectivity of the data samples on the graph, but also by adjusting the size of the clusters with respect to the local geometry, it provides a solution for both of the problems described in Figures 1(b) and 1(c).

In the proposed GOC method, contrary to AGNN, we need to define a strategy to select the PCA basis that best fits a given test patch. Given a test patch y_j , we propose to select a basis Φ_k by taking into account the distance between y_j and the centroid μ_k of the cluster S_k (corresponding to Φ_k), as well as the agreement between y_j and the principal directions in Φ_k . Let $\Phi_k^r = [\phi_1 \dots \phi_r]$ denote the submatrix of Φ_k consisting of the first r principal vectors, which give the directions that determine the main orientation of the cluster. We then choose

the basis Φ_k that minimizes

$$k = \arg \min_{k'} \left\{ \|y_j - \mu_{k'}\|_2 - \gamma \left\| (\Phi_k^r)^T \frac{y_j - \mu_{k'}}{\|y_j - \mu_{k'}\|_2} \right\|_2 \right\} \quad (17)$$

where $\gamma > 0$ is a weight parameter. While the first term above minimizes the distance to the centroid of the cluster, the second term maximizes the correlation between the relative patch position $y_j - \mu_{k'}$ and the most significant principal directions. Once the basis index k is determined as above, the test patch y_j is reconstructed based on a sparse representation in Φ_k .

VI. EXPERIMENTS

In this section, we experimentally evaluate the performance of the proposed methods. In Section VI-A we first present an experiment which demonstrates that the proposed neighborhood selection strategies can be successfully used for capturing structural similarities of images. Then, in Section VI-B we test our algorithms in image super-resolution. Finally in Section VI-C we discuss the applicability of the proposed methods to other image restoration applications and overview some results obtained in image deblurring and denoising.

A. Transformation-invariant patch similarity analysis

Natural images often contain different observations of the same structure in different regions of the image. Patches that share a common structure may be generated from the same reference pattern with respect to a transformation model that can possibly be parameterized with a few parameters. One example to parametrizable transformation models is geometric transformations. In this section, we evaluate the performance of the proposed AGNN strategy in capturing structural similarities between image patches in a transformation-invariant way. We generate a collection of patches of size 10×10 pixels, by taking a small set of reference patches and applying geometric transformations consisting of a rotation with different angles to each reference patch to obtain a set of geometrically transformed versions of it. Figure 4 shows two reference patches and some of their rotated versions. The data set used in the experiment is generated from 10 reference patches, which are rotated at intervals of 5 degrees.

In order to evaluate the performance of transformation-invariant similarity assessment, we look for the nearest neighbors of each patch in the whole collection and identify the “correct” neighbors as the ones sharing the same structure, i.e., the patches generated from the same reference patch. Three nearest neighbor selection strategies are tested in the experiment, which are AGNN, neighbor selection with respect to Euclidean distance, and K-means clustering. In AGNN, the neighborhood size that gives the best algorithm performance is used. The Euclidean distance uses the same neighborhood size as AGNN, and the number of clusters in K-means is set as the true number of clusters, i.e., the number of reference patches generating the data set. The correct clustering rates are shown in Figure 5, which are the percentage of patches that are correctly present in a cluster (each neighborhood is considered as a cluster in AGNN and Euclidean distance). The horizontal axis shows the number of clusters (i.e., number of reference

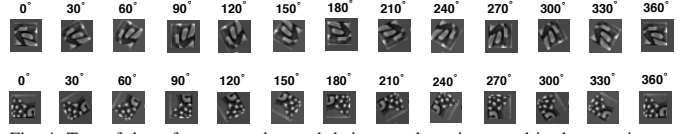


Fig. 4: Two of the reference patches and their rotated versions used in the experiment

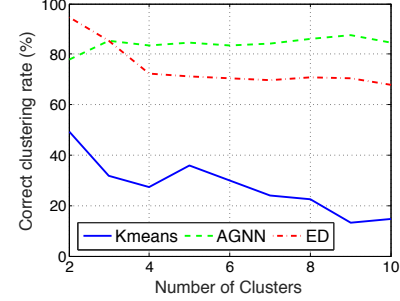


Fig. 5: Percentage of patches correctly included in the clusters

patches) used in different repetitions of the experiment. It can be observed that the AGNN method yields the best transformation-invariant similarity assessment performance. Contrary to methods based on simple Euclidean distance, AGNN measures the similarity of two patches by tracing all paths on the manifold joining them. Therefore, it is capable of following the gradual transformations of structures on the patch manifold and thus identifying structural similarities of patches in a transformation-invariant manner.

B. Image super-resolution

In this section, we demonstrate the benefits of our neighborhood selection strategies in the context of the NCSR algorithm [1], which leads to state-of-the-art performance in image super-resolution.

The NCSR algorithm [1] is an image restoration method that reconstructs image patches by selecting a model among a set of local PCA bases. This strategy exploits the image nonlocal self-similarity to obtain estimates of the sparse coding coefficients of the observed image. The method first clusters training patches with the K-means algorithm and then adopts the adaptive sparse domain selection strategy proposed in [2] to learn a local PCA basis for each cluster from the estimated high-resolution (HR) images. After the patches are coded, the NCSR objective function is optimized with the Iterative Shrinkage Thresholding (IST) algorithm proposed in [35]. The clustering of training patches with the K-means algorithm in [1] is based on adopting the Euclidean distance as a dissimilarity measure. The purpose of our experiments is then to show that the proposed geometry-based nearest neighbor selection methods can be used for improving the performance of an image super-resolution algorithm such as NCSR.

We now describe the details of our experimental setting for the super-resolution problem. In the inverse problem $y = \Theta x + \nu$ in (1), x and y denote respectively the lexicographical representations of the unknown image X and the degraded image Y . The degradation matrix $\Theta = DH$ is composed of a down-sampling operator D with a scale factor of $q = 3$ and a

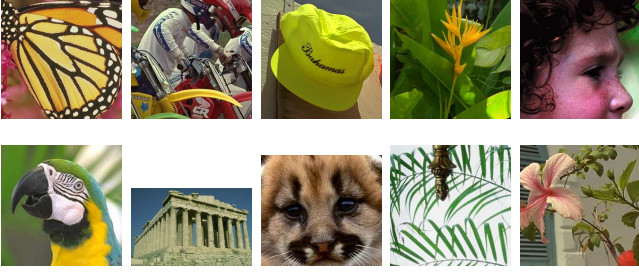


Fig. 6: Test images for super-resolution: Butterfly, Bike, Hat, Plants, Girl, Parrot, Parthenon, Raccoon, Leaves, Flower.

Gaussian filter H of size 7×7 with a standard deviation of 1.6, and ν is an additive noise. We aim to recover the unknown image vector \mathbf{x} from the observed image vector \mathbf{y} . We evaluate the proposed algorithms on the 10 images presented in Figure 6, which differ in their frequency characteristics and content. For color images, we apply the single image SR algorithm only on the luminance channel and we compute the PSNR and SSIM [36] only on the luminance channel for coherence. Besides PSNR and SSIM, the visual quality of the images is also used as a comparison metric.

Overlapping patches of size 6×6 are used in the experiments. The original NCSR algorithm initializes the training set \mathcal{D} by extracting patches from several images in the scale space of the HR image. However, in our implementation we initialize the set of training patches by extracting them only from the low-resolution image; i.e., the m initial training patches $d_i \in \mathbb{R}^n$ in $\mathcal{D} = \{d_i\}_{i=1}^m$ are extracted from the observed low-resolution (LR) image vector \mathbf{y} . We learn online PCA bases using the training patches in \mathcal{D} with the proposed AGNN and GOC methods. In the original NCSR method, in every P iterations of the IST algorithm, the training set \mathcal{D} is updated by extracting the training patches from the current version of the reconstructed image $\hat{\mathbf{x}}$ and the PCA bases are updated as well by repeating the neighborhood selection with the updated training data. In our experiments, we use the same training patches \mathcal{D} for the whole algorithm.

In Section VI-B1, we evaluate our methods AGNN and GOC by comparing their performance to some other clustering or nearest neighbor selection strategies in super-resolution. In Section VI-B2, we provide comparative experiments with several widely used super-resolution algorithms and show that our proposed manifold-based neighborhood selection techniques can be used for improving the state of the art in super-resolution.

In Sections VI-B1 and VI-B2, the results for the algorithms in comparison are obtained by using the software packages made publicly available by the corresponding authors ¹.

1) *Performance Evaluation of AGNN and GOC:* We compare the proposed AGNN and GOC methods with 4 different clustering algorithms; namely, the K-means algorithm (Kmeans), Fuzzy C-means clustering algorithm (FCM) [32], Spectral Clustering (SC) [20], Replicator Graph Clustering (RGC) [18]; and also with K-NN search using geodesic

distance (GeoD). Among the clustering methods, Kmeans and FCM employ the Euclidean distance as a dissimilarity measure, while SC and RGC are graph-based methods that consider the manifold structure of data. When testing these four methods, we cluster the training patches and compute a PCA basis for each cluster. Then, given a test patch, the basis of the cluster whose centroid has the smallest distance to the test patch is selected as done in the original NCSR algorithm where K-means is used. In the GeoD method, each test patch is reconstructed with the PCA basis computed from its nearest neighbors with respect to the geodesic distance numerically computed with Dijkstra's algorithm [37]. The idea of nearest neighbor selection with respect to the geodesic distance is also in the core of the methods proposed in [27] and [28]. Note that the four reference clustering methods and GOC provide nonadaptive solutions for training subset selection, while the GeoD and the AGNN methods are adaptive.

The parameters of the AGNN algorithm are set as $s = 35$ (number of nearest neighbors in the diffusion stage of RGC [18]), $\kappa = 2$ (number of iterations for diffusing the affinity matrix), $c_1 = 10$ (Gaussian kernel scale), and $c_2 = 0.9$ (affinity threshold). The parameters of the GOC algorithm are set as $C = 64$ (number of clusters), $c_3 = 0.5$ (threshold defining the decay function), $\gamma = 150$, and $r = 8$ (parameters for selecting a PCA basis for each test patch). The number of clusters in the other four clustering methods in comparison are also set to the same value as $C = 64$. The size of the clusters with the FCM algorithm are selected roughly the same as the cluster sizes computed with K-means. The total number of iterations and the number of PCA basis updates are chosen as 1000 and 4 in the NCSR algorithm. All the general parameters for the NCSR algorithm are selected as Dong et al. [1]. In this way, we can maintain consistency in the comparison of the methods related to NCSR algorithm.

We evaluate the GOC algorithm in three different settings. In the first setting the cluster size parameters L and K are estimated adaptively for each cluster with the strategy proposed in Algorithm 2, which is denoted as aGOC. In the second setting, denoted avGOC, the parameters L and K are not adapted to each cluster; all clusters are formed with the same parameter values, where L and K are computed by minimizing the average value of coefficient decay function $\tilde{I}(L, K)$ over all clusters of the same image. The parameters are thus adapted to the images, but not to the individual clusters of patches of an image. Finally, in the third setting, denoted mGOC, the parameters L and K are manually entered and used for all clusters of the same image. The parameter values provided to the algorithm for each image are set as the best values obtained with an exhaustive search. Therefore, mGOC can be considered as an oracle setting.

The results are presented in Figure 7, Figure 8, and Table I. Figures 7 and 8 provide a visual comparison between the image reconstruction qualities obtained with the K-means clustering algorithm and the proposed AGNN and GOC methods for the Butterfly and the Hat images. It is observed that AGNN and GOC produce sharper edges than K-means. Moreover, the visual artifacts produced by K-means such as the phantom perpendicular bands on the black stripes of the

¹We would like to thank the authors of [1], [2], [18], [19], [20], [32] and [37] for making their software packages publicly available.

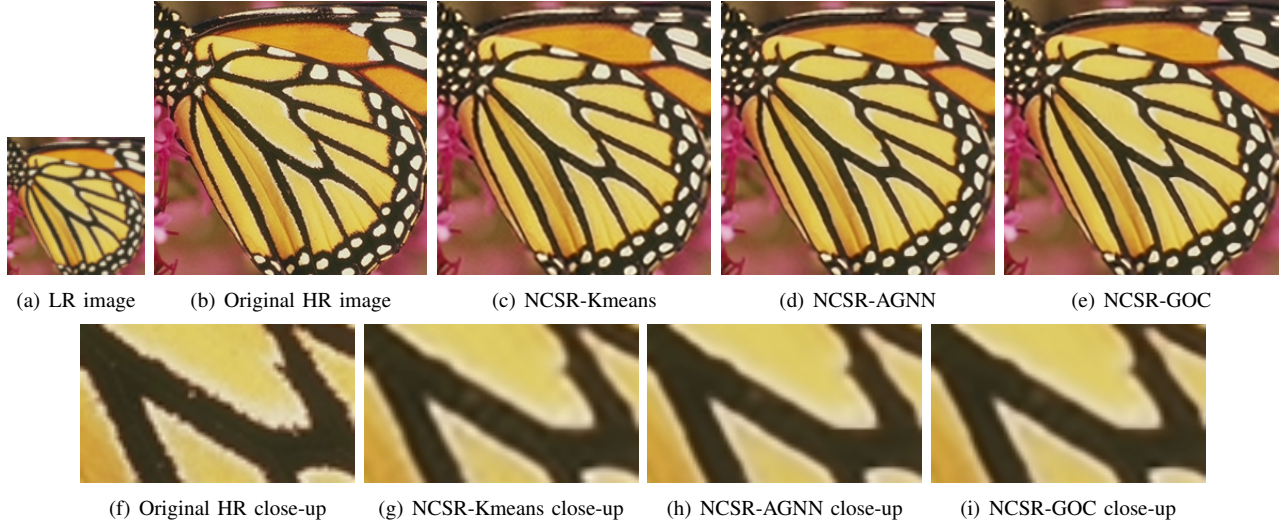


Fig. 7: Comparison of SR results ($\times 3$). It can be observed that NCSR-AGNN and NCSR-GOC reconstruct edges with a higher contrast than NCSR-Kmeans. Artifacts visible with NCSR-Kmeans (e.g., the oscillatory phantom bands perpendicular to the black stripes on the butterfly's wing) are significantly reduced with NCSR-AGNN and NCSR-GOC.

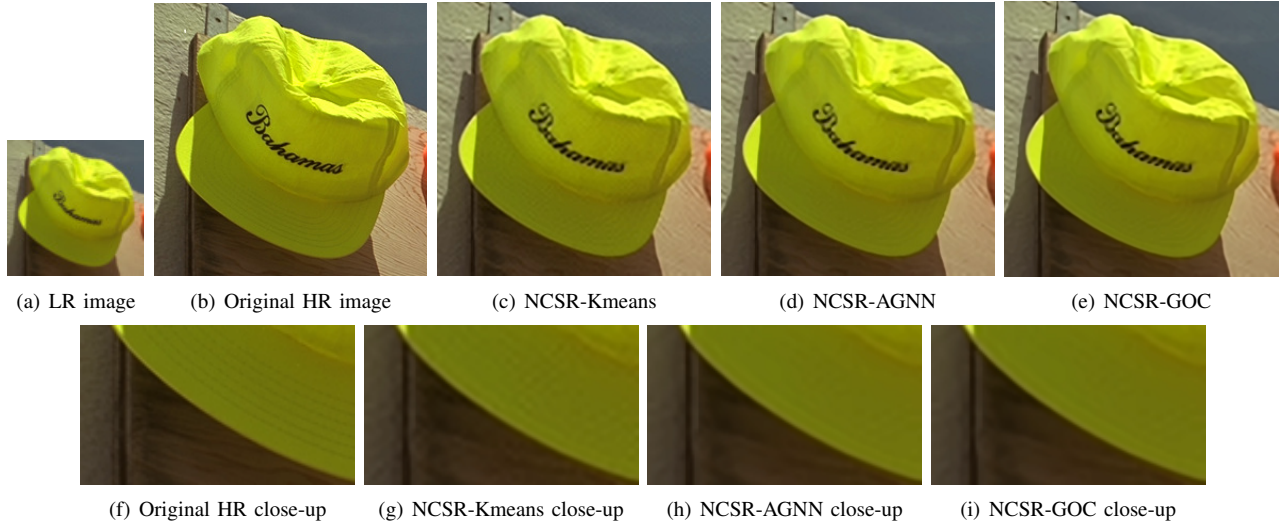


Fig. 8: Comparison of SR results ($\times 3$). NCSR-Kmeans produces artifacts such as the checkerboard-like noise patterns visible on plain regions of the cap, which are prevented by NCSR-AGNN or NCSR-GOC.

butterfly and the checkerboard-like noise patterns on the cap are significantly reduced with AGNN and GOC. The efficiency of the proposed methods for removing these artifacts can be explained as follows. When image patches are clustered with the K-means algorithm, the similarity between patches is measured with the Euclidean distance. Therefore, when reconstructing a test patch, the algorithm tends to use a basis computed with patches that have similar intensity values. The nonuniformity of the pixel intensities along the black stripes of the LR Butterfly image thus propagates to the reconstructed HR image as well, which produces the phantom bands on the wing (due to the too low resolution, the black stripes on the LR image contain periodically appearing clear pixels contaminated by the yellow plain regions on the wing). Similarly, in the Hat image, the clusters used in learning a basis for reconstructing the plain regions on the cap contain also patches extracted from the wall, which have a similar intensity with the cap.

This reproduces the shadowy patterns of the wall also on the cap. On the other hand, the AGNN method groups together patches that have a connection on the data graph. As the patches are extracted with overlapping windows shifting by one pixel, AGNN and GOC may have a stronger tendency than K-means for favoring patches from nearby or similar regions on the image that all share a common structure, which is also confirmed by the experiment in Section VI-A. The proposed methods yield local bases better fitted to the characteristics of patches, therefore, less artifacts are observed.

In Table I the performance of the compared clustering methods are measured with the PSNR and the SSIM metrics. Graph-based methods are generally seen to yield a better performance than methods based on Euclidean distance. This confirms the intuition that motivates our study; when selecting neighborhoods for learning local models, the geometry of the data should be respected. As far as the average performance is

TABLE I: PSNR (top row, in dB) and SSIM (bottom row) results for the luminance components of super-resolved HR images for different clustering or neighborhood selection approaches: Spectral Clustering (SC) [20]; Fuzzy C-means clustering algorithm (FCM) [32]; K-means clustering (Kmeans); Replicator Graph Clustering (RGC) [18]; kNN search with Dijkstra Algorithm (GeoD) [37]; and our methods GOC and AGNN.

Images	Butterfly	Bike	Hat	Plants	Girl	Parrot	Parthenon	Raccoon	Leaves	Flower	Average
SC [20]	28.15	24.73	31.28	33.98	33.65	30.45	27.19	29.24	27.50	29.45	29.56
	0.9193	0.8026	0.8723	0.9198	0.8255	0.9170	0.7509	0.7659	0.9242	0.8567	0.8554
FCM [32]	28.20	24.76	31.25	33.99	33.65	30.47	27.25	29.25	27.68	29.50	29.60
	0.9205	0.8040	0.8726	0.9205	0.8256	0.9174	0.7531	0.7663	0.9271	0.8575	0.8565
Kmeans	28.14	24.79	31.31	34.07	33.64	30.53	27.20	29.28	27.67	29.47	29.61
	0.9204	0.8050	0.8730	0.9213	0.8254	0.9178	0.7517	0.7668	0.9265	0.8567	0.8565
RGC [18]	28.45	24.80	31.37	34.20	33.65	30.57	27.22	29.27	27.90	29.50	29.69
	0.9234	0.8061	0.8739	0.9219	0.8254	0.9181	0.7525	0.7658	0.9317	0.8576	0.8576
GeoD [37]	28.61	24.82	31.42	34.16	33.63	30.44	27.24	29.25	27.98	29.54	29.71
	0.9257	0.8070	0.8746	0.9219	0.8250	0.9178	0.7530	0.7650	0.9323	0.8587	0.8581
avGOC	28.34	24.85	31.42	34.17	33.66	30.68	27.23	29.28	27.89	29.55	29.71
	0.9222	0.8076	0.8747	0.9224	0.8258	0.9191	0.7528	0.7668	0.9317	0.8591	0.8582
aGOC	28.46	24.85	31.44	34.18	33.65	30.63	27.23	29.27	27.92	29.54	29.72
	0.9239	0.8082	0.8744	0.9227	0.8257	0.9187	0.7530	0.7663	0.9324	0.8588	0.8584
mGOC	28.54	24.90	31.43	34.20	33.67	30.71	27.25	29.28	27.95	29.55	29.75
	0.9251	0.8085	0.8748	0.9222	0.8261	0.9192	0.7530	0.7671	0.9324	0.8593	0.8588
AGNN	28.78	24.87	31.46	34.16	33.67	30.60	27.29	29.26	28.01	29.61	29.77
	0.9266	0.8081	0.8749	0.9218	0.8260	0.9188	0.7540	0.7661	0.9324	0.8601	0.8589

concerned, the AGNN method gives the highest reconstruction quality and is followed by the GOC method. The performance difference between AGNN and GOC can be justified with the fact that the training subset selection is adaptive to the test patches in AGNN, while GOC is a nonadaptive method that offers a less complex solution. In particular, with a non-optimized implementation of our algorithms, we have observed that GOC has roughly the same computation time as K-means, while the computation time of AGNN is around three times K-means and GOC in the tested images on an Intel Core i5 2.6GHz under the Matlab R2015a programming environment, as shown in Table III. After the proposed AGNN and GOC methods, GeoD gives the best average performance. While this adaptive method ensures a good reconstruction quality, it requires the computation of the geodesic distance between each test patch and all training patches. Therefore, it is computationally very complex. Although several works such as [27] and [28] provide solutions for fast approximations of the geodesic distance, we observe that in terms of reconstruction quality AGNN performs better than GeoD in most images. This suggests that using a globally consistent affinity measure optimized with respect to the entire graph topology provides a more refined and precise similarity metric than the geodesic distance, which only takes into account the shortest paths between samples.

Concerning the performances of the clustering methods on the individual images, an important conclusion is that geometry-based methods yield a better performance especially for images that contain patches of rich texture. The AGNN and GOC methods provide a performance gain of respectively 0.64 dB and 0.4 dB over K-means (used in the original NCSR method) for the Butterfly image. Meanwhile, all clustering methods give similar reconstruction qualities for the Girl

image. This discrepancy can be explained with the difference in the characteristics of the patch manifolds of these two images. The patches of the Butterfly image contain high-frequency textures; therefore, the patch manifold has a large curvature (see, e.g., [38] for a study of the relation between the manifold curvature and the image characteristics). Consequently, the proposed methods adapted to the local geometry of the manifold perform better on this image. On the other hand, the Girl image mostly contains weakly textured low-frequency patches, which generate a rather flat patch manifold of small curvature. The Euclidean distance is more reliable as a dissimilarity measure on flat manifolds compared to curved manifolds as it gets closer to the geodesic distance. Hence, the performance gain of geometry-based methods over K-means is much smaller on the Girl image compared to Butterfly.

Next, the comparison of the three modes of the GOC algorithm shows that aGOC and avGOC yield reconstruction qualities that are close to that of the oracle method mGOC. This suggests that setting the parameters L and K with respect to the PCA coefficient decay rates as proposed in Algorithm 2 provides an efficient strategy for the automatic determination of cluster sizes. While the average performances of aGOC and avGOC are quite close, interestingly, aGOC performs better than avGOC on Butterfly and Leaves. Both of these two images contain patches of quite varying characteristics, e.g., highly textured regions formed by repetitive edges as well as weakly textured regions. As the structures of the patches change significantly among different clusters in these images, optimizing the cluster size parameters individually for each cluster in aGOC has an advantage over using common parameters in avGOC.

2) *Improvements over the State of the Art in super-resolution:* In this section, we present an experimental com-

TABLE II: PSNR (top row, in dB) and SSIM (bottom row) results for the luminance components of super-resolved HR images for different super-resolution algorithms: Bicubic Interpolation; SPSR (Peleg et al.) [19]; ASDS (Dong et al.) [2]; NCSR (Dong et al.) [1]; NCSR with proposed GOC; NCSR with proposed AGNN.

Images	Butterfly	Bike	Hat	Plants	Leaves	Average	Parrot	Parthenon	Raccoon	Girl	Flower	Average
Bicubic	22.41	21.77	28.22	29.69	21.73	24.76	26.54	25.20	27.54	31.65	26.16	27.42
	0.7705	0.6299	0.8056	0.8286	0.7302	0.7530	0.8493	0.6528	0.6737	0.7671	0.7295	0.7345
SPSR [19]	26.74	24.31	30.84	32.83	25.84	28.11	29.68	26.77	29.00	33.40	28.89	29.55
	0.8973	0.7830	0.8674	0.9036	0.8892	0.8681	0.9089	0.7310	0.7562	0.8211	0.8415	0.8117
ASDS [2]	27.34	24.62	30.93	33.47	26.80	28.63	30.00	26.83	29.24	33.53	29.19	29.76
	0.9047	0.7962	0.8706	0.9095	0.9058	0.8774	0.9093	0.7349	0.7677	0.8242	0.8480	0.8168
NCSR [1]	28.07	24.74	31.29	34.05	27.46	29.12	30.49	27.18	29.27	33.66	29.50	30.02
	0.9156	0.8031	0.8704	0.9188	0.9219	0.8860	0.9147	0.7510	0.7707	0.8276	0.8563	0.8241
NCSR-GOC	28.47	24.85	31.44	34.16	28.05	29.39	30.71	27.23	29.28	33.65	29.58	30.09
	0.9241	0.8084	0.8747	0.9232	0.9339	0.8929	0.9192	0.7526	0.7666	0.8257	0.8600	0.8248
NCSR-AGNN	28.81	24.86	31.47	34.19	28.06	29.48	30.60	27.30	29.27	33.67	29.60	30.09
	0.9273	0.8080	0.8755	0.9223	0.9332	0.8933	0.9189	0.7546	0.7662	0.8261	0.8601	0.8252

TABLE III: Time (s) results for the luminance components of super-resolved HR images for NCSR (Dong et al.) [1]; NCSR with proposed GOC; NCSR with proposed AGNN.

Images	Butterfly	Bike	Hat	Plants	Leaves	Parrot	Parthenon	Raccoon	Girl	Flower	Average
NCSR [1]	261	229	213	229	233	220	481	362	213	226	267
NCSR-GOC	271	266	253	261	278	256	518	383	246	264	299
NCSR-AGNN	960	1039	467	578	1146	505	2541	1637	416	830	1012

parison of several popular super-resolution algorithms; namely, the bicubic interpolation algorithm, ASDS [2], SPSR [19], and NCSR [1]. We evaluate the performance of the NCSR algorithm under three different settings where the local bases are computed with K-means, AGNN, and GOC. The GOC method is used as in Algorithm 2 (denoted as aGOC in the previous experiments).

The experiments are conducted on the same images as in the previous set of experiments. The total number of iterations and the number of PCA basis updates of NCSR are selected respectively as 960 and 6, while the other parameters are chosen as before. The results presented in Table II show that the state of the art in super-resolution is led by the NCSR method [1]. The performance of NCSR is improved when it is coupled with the AGNN and GOC strategies for selecting local models. In Table II the images are divided into two categories as those with high-frequency and low-frequency content. The average PSNR and SSIM metrics are reported in both groups. It can be observed that the advantage of the proposed neighborhood selection strategies over K-means is especially significant for high-frequency images. In images with low-frequency content, K-means gives the same performance as the proposed methods. As the patch manifold gets flatter, clusters obtained with K-means and the proposed methods get similar. Hence, we may conclude that the proposed geometry-based neighborhood selection methods can be successfully used for improving the state of the art in image super-resolution, whose efficacy is especially observable for sharp images rich in high-frequency texture.

C. Applicability to other restoration problems

The methods proposed in this paper aim to determine patch similarity in a geometry-aware manner and can potentially be

used in other image restoration applications than superresolution as well. We now briefly overview the applicability of our algorithms to other problems.

We first evaluate our method in the image deblurring application. We compare GOC with the K-means clustering algorithm within the framework of the NCSR method [1]. The algorithms are tested on the images shown in Figure 9. A Gaussian blur kernel of standard deviation 1.6 pixels is used. Along with the blurring, the images are also corrupted with an additive white Gaussian noise of standard deviation $\sqrt{2}$. The parameters of GOC are set as $C = 64$ (number of clusters), $c_3 = 0.5$ (threshold defining the decay function), $\gamma = 150$, and $r = 8$ (parameters for selecting a PCA basis for each test patch). All the general parameters for the NCSR algorithm are selected as Dong et al. [1] in order to maintain the consistency.

The PSNR and FSIM [39] measures of the reconstruction qualities are presented in Table IV. The results obtained with the image restoration algorithms FISTA (Portilla et al.) [40] and ASDS (Dong et al.) [2] reported in [1] for the same experiments are also given for the purpose of comparison. The results show that the proposed GOC algorithm can be effectively used for improving the image reconstruction quality of the NCSR method in deblurring applications. The GOC method either outperforms the K-means clustering algorithm or yields a quite close performance when coupled with NCSR. Moreover, one can observe that the best average PSNR value is given by the proposed method, whose benefits are especially observable for images with significant high-frequency components such as Butterfly, Cameraman, and Leaves. More detailed results on image deblurring can be found in the accompanying technical report [41], where the algorithms are also tested with a uniform blur kernel and their performance comparison is observed to be similar.

TABLE IV: PSNR (top row, in dB) and FSIM (bottom row) results for the luminance components of deblurred images for different deblurring algorithms for a Gaussian blur kernel of standard deviation 1.6 pixels and an additive white Gaussian noise of standard deviation $\sqrt{2}$: FISTA (Portilla et al.) [40]; ASDS (Dong et al.) [2]; NCSR (Dong et al.) [1]; NCSR with proposed GOC.

Images	Butterfly	Boats	C. Man	House	Parrot	Lena	Barbara	Starfish	Peppers	Leaves	Average
FISTA [40]	30.36	29.36	26.81	31.50	31.23	29.47	25.03	29.65	29.42	29.36	29.22
	0.9452	0.9024	0.8845	0.8968	0.9290	0.9011	0.8415	0.9256	0.9057	0.9393	0.9071
ASDS [2]	29.83	30.27	27.29	31.87	32.93	30.36	27.05	31.91	28.95	30.62	30.11
	0.9126	0.9064	0.8637	0.8978	0.9576	0.9058	0.8881	0.9491	0.9039	0.9304	0.9115
NCSR [1]	30.84	31.37	28.27	33.69	33.40	31.17	28.02	32.23	30.01	31.62	31.06
	0.9379	0.9348	0.9044	0.9339	0.9589	0.9360	0.9108	0.9533	0.9300	0.9514	0.9351
NCSR-GOC	31.32	31.48	28.44	33.80	33.45	31.28	27.45	32.27	30.27	32.04	31.18
	0.9486	0.9413	0.9153	0.9375	0.9594	0.9429	0.9014	0.9554	0.9389	0.9587	0.9399



Fig. 9: Test images for deblurring: Butterfly, Boats, Cameraman, House, Parrot, Lena, Barbara, Starfish, Peppers, Leaves.

Next, we discuss the applicability of our methods to image denoising. In the technical report [41], we present experimental results where we test our AGNN algorithm in image denoising. We refer the reader to [41] for a detailed discussion of the results and briefly overview our findings here. The AGNN algorithm is compared to K-means for the clustering of the image patches in image denoising with the NCSR method [1]. The two versions of the NCSR algorithm, coupled with K-means and AGNN, are also compared with several other denoising methods on a set of images corrupted with Gaussian noise of standard deviation $\sigma = [5 \ 10 \ 15 \ 20 \ 50 \ 100]$. While the overall performances of the compared algorithms are close, SAPCA-BM3D [42] usually gives the best denoising performance. The proposed NCSR-AGNN algorithm yields a very similar performance to NCSR. A very slight improvement in average PSNR is obtained over NCSR at small noise levels (smaller than $\sigma = 20$), while this small advantage is lost at larger noise levels. The results show that the performance of NCSR-AGNN tends to be better than NCSR on images with strong and oscillatory high-frequency textures. This may be due to the particular geometry of the patch manifold in such images, which is easier to identify under noise, so that the consideration of the geometry in assigning the similarities helps improve the denoising performance.

VII. CONCLUSION

In this paper, we have focused on the problem of selecting local subsets of training data samples that can be used for learning local models for image super-resolution. This study has been motivated by the observation that the Euclidean distance may not always be a good dissimilarity measure

for comparing data samples lying on a manifold. We have proposed two methods for data subset selection by taking into account the geometry of the data, which is assumed to lie on a manifold. Although the addressed problem has close links with manifold clustering, it differs by the fact that the goal here is not to obtain a partitioning of data, but instead select a local subset of training data that can be used for learning a good model for sparse reconstruction of a given input test sample. The performance of the methods has been demonstrated in a super-resolution application leading to a novel single-image super-resolution algorithm which outperforms reference methods. The applicability to other restoration problems has also been discussed.

ACKNOWLEDGMENT

The first author would like to thank CAPES Brazilian agency for the financial support (PDSE scholarship 18385-12-5). The authors also thank J  r  my Aghaei Mazaheri for the helpful discussions.

REFERENCES

- [1] W. Dong, L. Zhang, G. Shi, and X. Li, "Nonlocally centralized sparse representation for image restoration," *IEEE Transactions on Image Processing*, vol. 22, no. 4, pp. 1620–1630, apr 2013.
- [2] W. Dong, L. Zhang, G. Shi, and X. Wu, "Image deblurring and super-resolution by adaptive sparse domain selection and adaptive regularization," *IEEE Transactions on Image Processing*, vol. 20, no. 7, pp. 1838–1857, jul 2011.
- [3] J. Yang, J. Wright, T. Huang, and Y. Ma, "Image Super-Resolution as Sparse Representation of Raw Image Patches," in *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–8.
- [4] J. Yang, J. Wright, T. S. Huang, and Yi Ma, "Image Super-Resolution Via Sparse Representation," *IEEE Transactions on Image Processing*, vol. 19, no. 11, pp. 2861–2873, nov 2010.
- [5] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Transactions on Image Processing*, vol. 15, no. 12, pp. 3736–45, dec 2006.
- [6] W. Dong, X. Li, L. Zhang, and G. Shi, "Sparsity-based image denoising via dictionary learning and structural clustering," in *Proceedings IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR)*, jun 2011, pp. 457–464.
- [7] D. L. Donoho, "Compressed sensing," *IEEE Transactions on Information Theory*, vol. 52, pp. 1289–1306, 2006.
- [8] E. J. Candes and T. Tao, "Near-Optimal Signal Recovery From Random Projections: Universal Encoding Strategies?" *IEEE Transactions on Information Theory*, vol. 52, no. 12, pp. 5406–5425, dec 2006.
- [9] E. J. Cand  s, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Transactions on Information Theory*, vol. 52, no. 2, pp. 489–509, 2006.

- [10] J. Ni, P. Turaga, V. M. Patel, and R. Chellappa, "Example-driven manifold priors for image deconvolution," *IEEE Transactions on Image Processing*, vol. 20, no. 11, pp. 3086–3096, nov 2011.
- [11] J. Salmon, Z. Harmany, C.-A. Deledalle, and R. Willett, "Poisson Noise Reduction with Non-local PCA," *Journal of Mathematical Imaging and Vision*, vol. 48, no. 2, pp. 279–294, feb 2014.
- [12] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," in *Proceedings Workshop on Signal Processing with Adaptive Sparse Structured Representations (SPARS)*, feb 2006, pp. 606 414–606 414–12.
- [13] A. Danielyan, A. Foi, V. Katkovnik, and K. Egiazarian, "Denoising of multispectral images via nonlocal groupwise spectrum-PCA," in *Conference on Colour in Graphics, Imaging, and Vision*, vol. 2010, no. 1, 2010, pp. 261–266.
- [14] A. B. Lee, K. S. Pedersen, and D. Mumford, "The Nonlinear Statistics of High-Contrast Patches in Natural Images," *International Journal of Computer Vision*, vol. 54, pp. 83–103, 2003.
- [15] G. Peyré, "Manifold models for signals and images," *Computer Vision and Image Understanding*, vol. 113, no. September 2008, pp. 249–260, 2009.
- [16] D. N. Kaslovsky and F. G. Meyer, "Overcoming noise, avoiding curvature: Optimal scale selection for tangent plane recovery," in *Proceedings IEEE Statistical Signal Processing Workshop (SSP)*, 2012, pp. 892–895.
- [17] H. Tyagi, E. Vural, and P. Frossard, "Tangent space estimation for smooth embeddings of Riemannian manifolds," *Information and Inference*, vol. 2, pp. 69–114, 2013.
- [18] M. Donoser, "Replicator Graph Clustering," in *Proceedings of the British Machine Vision Conference (BMVC)*, 2013, pp. 38.1–38.11.
- [19] T. Peleg and M. Elad, "A statistical prediction model based on sparse representations for single image super-resolution," *IEEE Transactions on Image Processing*, vol. 23, no. 6, pp. 2569–2582, jun 2014.
- [20] J. Shi and J. Malik, "Normalized cuts and image segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, 2000.
- [21] A. Y. Ng, M. I. Jordan, and Y. Weiss, "On Spectral Clustering: Analysis and an algorithm," in *Proceedings Advances in Neural Information Processing Systems*, 2001, pp. 849–856.
- [22] M. Belkin and P. Niyogi, "Laplacian Eigenmaps for Dimensionality Reduction and Data Representation," *Neural Computation*, vol. 15, pp. 1373–1396, 2003.
- [23] N. Asgharbeygi and A. Maleki, "Geodesic K-means clustering," in *Proceedings International Conference on Pattern Recognition (ICPR)*, Tampa, 2008, pp. 1–4.
- [24] E. Tu, L. Cao, J. Yang, and N. Kasabov, "A novel graph-based k-means for nonlinear manifold clustering and representative selection," *Neurocomputing*, vol. 143, pp. 1–14, 2014.
- [25] M. Breitenbach and G. Z. Grudic, "Clustering through ranking on manifolds," in *Proceedings of the International Conference on Machine learning (ICML)*, 2005, pp. 73–80.
- [26] D. Zhou, O. Bousquet, T. N. Lal, J. Weston, and B. Schölkopf, "Learning with Local and Global Consistency," in *Proceedings Advances in Neural Information Processing Systems 16 (NIPS)*, 2004, pp. 321–328.
- [27] P. Turaga and R. Chellappa, "Nearest-Neighbor Search Algorithms on Non-Euclidean Manifolds for Computer Vision Applications," in *Proceedings of the Indian Conference on Computer Vision, Graphics and Image Processing*, New York, New York, USA, 2010, pp. 282–289.
- [28] R. Chaudhry and Y. Ivanov, "Fast Approximate Nearest Neighbor Methods for Non-Euclidean Manifolds with Applications to Human Activity Analysis in Videos," in *Proceedings European Conference on Computer Vision (ECCV)*, vol. 8689, Heraklion, 2010, pp. 735–748.
- [29] R. Souvenir and R. Piess, "Manifold clustering," in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, vol. I, 2005, pp. 648–653.
- [30] E. Elhamifar and R. Vidal, "Sparse Manifold Clustering and Embedding," in *Proceedings Advances in Neural Information Processing Systems 24 (Nips)*, 2011, pp. 55–63.
- [31] A. Goh and R. Vidal, "Clustering and dimensionality reduction on Riemannian manifolds," in *Proceedings IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2008, pp. 1–7.
- [32] J. C. Bezdek, R. Ehrlich, and W. Full, "FCM: The fuzzy c-means clustering algorithm," *Computers & Geosciences*, vol. 10, no. 2, pp. 191–203, 1984.
- [33] J. Kim, K. H. Shim, and S. Choi, "Soft geodesic kernel K-means," in *Proceedings IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, vol. 2, no. 3, 2007, pp. 429–432.
- [34] G. Yu, G. Sapiro, and S. Mallat, "Solving inverse problems with piecewise linear estimators: From gaussian mixture models to structured sparsity," *IEEE Transactions on Image Processing*, vol. 21, no. 5, pp. 2481–2499, may 2012.
- [35] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with a sparsity constraint," *Communications on Pure and Applied Mathematics*, vol. 57, no. 11, pp. 1413–1457, nov 2004.
- [36] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, apr 2004.
- [37] E. W. Dijkstra, "A note on two problems in connexion with graphs," *Numerische Mathematik*, vol. 1, pp. 269–271, 1959.
- [38] E. Vural and P. Frossard, "Curvature analysis of pattern transformation manifolds," in *Proceedings IEEE International Conference on Image Processing (ICIP)*, sep 2010, pp. 2689–2692.
- [39] L. Zhang, L. Zhang, X. Mou, and D. Zhang, "FSIM: A Feature Similarity Index for Image Quality Assessment," *IEEE Transactions on Image Processing*, vol. 20, no. 8, pp. 2378–2386, aug 2011.
- [40] J. Portilla, "Image restoration through L0 analysis-based sparse optimization in tight frames," in *Proceedings IEEE International Conference on Image Processing (ICIP)*, 2009, pp. 3909–3912.
- [41] J. C. Ferreira, E. Vural, and C. Guillemot, "Geometry-Aware Neighborhood Search for Learning Local Models for Image Reconstruction," may 2015. [Online]. Available: <http://arxiv.org/abs/1505.01429>
- [42] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, "From Local Kernel to Nonlocal Multiple-Model Image Denoising," *International Journal of Computer Vision*, vol. 86, no. 1, pp. 1–32, jan 2010.